

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

Patent No. 7,627,573

Confirmation No. 9536

Issued: December 1, 2009

Name of Patentee: Williams, Jr. et al.

Patent Title: Bypassing Native Storage Operations By Communicating Protected Data
Within Locking Messages Using A Lock Manager Independent Of The
Storage Mechanism

**REQUEST FOR CERTIFICATE OF CORRECTION OF
PATENT FOR APPLICANTS' MISTAKE (37 C.F.R. § 1.323)**

Attn: Certificate of Correction Branch
Commissioner for Patents
P.O. Box 1450
Alexandria, VA 22313-1450

Dear Sir:

It is requested that a Certificate of Correction be issued to correct typographical errors found the above-identified patent. Attached hereto is a Certificate of Correction which indicates the requested corrections.

In the issued patent, the spelling of the word "requester" was spelled as "requester" one hundred twenty-three (123) times and as "requestor" sixteen (16) times. In order to avoid any potential, yet unlikely, confusion, it is requested that the Office issue a Certificate of Correction to make the spelling of this word consistent throughout the issued Patent. Applicants consulted, by telephone, with Examiner Hicks, and he has no issue with these typographical errors being corrected. For your convenience, also attached are copies of selected pages from the issued patent with errors highlighted.

In re US Patent No. 7,627,573

A certificate of correction is proper as these errors are of the nature of an obvious typographical error, correction of this error does not constitute new matter, correction of this error does not require reexamination, and the requisite fee under 37 CFR 1.20(a) was paid via EFS-Web with the submission of this Request for Certificate of Correction. Finally, any additional fees that may be due in connection with this paper are authorized to be charged to Deposit Account No. 501430.

Respectfully submitted,
The Law Office of Kirk D. Williams

Date:

JAN 29, 2010 By 

Kirk D. Williams, Reg. No. 42,229
One of the Attorneys for Applicants
CUSTOMER NUMBER 26327
The Law Office of Kirk D. Williams
P.O. Box 39425, Denver, CO 80239-0425
303-282-0151 (telephone), 303-778-0748 (facsimile)

**UNITED STATES PATENT AND TRADEMARK OFFICE
CERTIFICATE OF CORRECTION**

Page 1 of 1

PATENT NO. : 7,627,573
APPLICATION NO. : 10/811,044
DATED : December 1, 2009
INVENTOR(S) : Williams, Jr. et al.

It is certified that an error appears or errors appear in the above-identified patent and that said Letters Patent is hereby corrected as shown below:

Face of the Patent, Abstract (57):

line 14, replace "requestor" with – requester –
line 16, replace "requestor" with – requester –

Col. 2, line 15, replace "requestor" with – requester –

Col. 3, line 15, replace "requestor" with – requester –

Col. 3, line 19, replace "requestor" with – requester –

Col. 6, line 42, replace "requestor" with – requester –

Col. 6, line 49, replace "requestor" with – requester –

Col. 7, line 47, replace "requestor" with – requester –

Col. 12, Claim 1, line 26, replace "requestors" with – requesters –

Col. 13, Claim 12, line 54, replace "requestor" with – requester –

Col. 13, Claim 14, line 63, replace "requestor" with – requester –

Col. 13, Claim 15, line 67, replace "requestor" with – requester –

Col. 14, Claim 19, line 40, replace "requestor" with – requester –

Col. 14, Claim 20, line 44, replace "requestor" with – requester –

Col. 15, Claim 24, line 13, replace "requestor" with – requester –

Col. 16, Claim 25, line 4, replace "requestor" with – requester –

MAILING ADDRESS OF SENDER:

Kirk D. Williams, Reg. No. 42,229
Customer No. 26327
The Law Office of Kirk D. Williams
P.O. Box 39425, Denver, CO 80239



US007627573B2

(12) **United States Patent**
Williams, Jr. et al.

(10) **Patent No.:** **US 7,627,573 B2**

(45) **Date of Patent:** Dec. 1, 2009

(54) **BYPASSING NATIVE STORAGE OPERATIONS BY COMMUNICATING PROTECTED DATA WITHIN LOCKING MESSAGES USING A LOCK MANAGER INDEPENDENT OF THE STORAGE MECHANISM**

(75) Inventors: **John J. Williams, Jr.**, Pleasanton, CA (US); **John Andrew Fingerhut**, Campbell, CA (US); **Jonathan Rosen**, Cary, NC (US)

(73) Assignee: **Cisco Technology, Inc.**, San Jose, CA
(US)

(*) Notice: Subject to any disclaimer, the term of this patent is extended or adjusted under 35 U.S.C. 154(b) by 824 days.

(21) Appl. No.: 10/811,044

(22) Filed: **Mar. 27, 2004**

(65) **Prior Publication Data**

US 2005/0216461 A1 Sep. 29, 2005

(51) **Int. Cl.**
G06F 7/00 (2006.01)
G06F 17/30 (2006.01)

(52) U.S. Cl. 707/8; 707/9; 707/10

(58) **Field of Classification Search** 707/8-10,
707/1

See application file for complete search history.

(56) **References Cited**

U.S. PATENT DOCUMENTS

5,151,988	A *	9/1992	Yamagishi	707/8
5,327,556	A *	7/1994	Mohan et al.	707/8
5,551,046	A *	8/1996	Mohan et al.	707/8
5,845,147	A	12/1998	Vishlitzky et al.	
6,253,273	B1	6/2001	Blumenau	
6,553,464	B1	4/2003	Kamvysselis et al.	
7,120,631	B1 *	10/2006	Vahalia et al.	707/8
2002/0184216	A1 *	12/2002	Chandrasekaran et al.	707/8

OTHER PUBLICATIONS

Yun et al. "An Efficient Locking Protocol for Home Based Lazy Release Consistency", Proceedings of the First IEEE/ACM International Symposium on Cluster Computing and The Grid, pp. 527-532; May 2001.*

Amiri et al. ("Highly Concurrent Shared Storage", Proceedings of the International Conference On Distributed Computing Systems, Taipei, Apr. 2000.*

Trancoso et al., "The Impact of Speeding up Critical Sections with Data Prefetching and Forwarding", 1996 International Conference on Parallel Processing, pp. III-79-III-86, 1996 IEEE.*

Coulouris et al., *Distributed Systems Concepts and Design*, Third Edition, Addison-Wesley, Ch. 12, pp. 465-514, 2001.

* cited by examiner

Primary Examiner—Neveen Abel Jalil

Assistant Examiner—Michael J Hicks

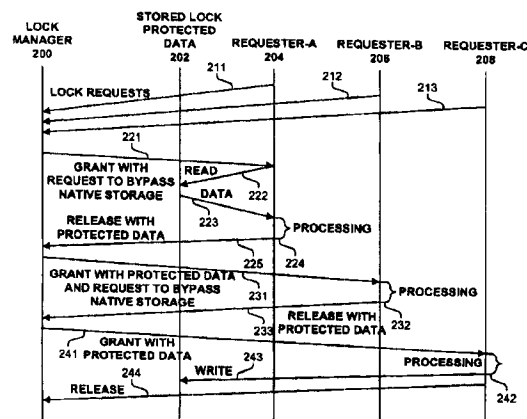
(74) *Attorney, Agent, or Firm*—The Law Office of Kirk D. Williams

(57) **ABSTRACT**

Data is protected using locks, with the protected data sometimes being included in the locking messages, which may reduce overall processing latency, and/or reduce a bandwidth requirement for and/or number of storage operations accessing the native storage of the protected data. For example, the lock manager receives lock requests from each of the requesters, and selectively grants the lock requests. The protected data is typically communicated in the locking messages when the lock is highly contested, or at least two request for access to the data are pending. The lock manager initiates the sequence by indicating in a grant message to a requester to include the protected data in its release message. The lock manager then copies this data received in the release message to its grant message to the next requestor. If no other requesters are waiting, the grant message includes an indication not to send the protected data, and thus the requestor typically stores this protected data to storage so it can be accessed in the future.

① ②
change to
requester "

26 Claims, 7 Drawing Sheets



1

**BYPASSING NATIVE STORAGE
OPERATIONS BY COMMUNICATING
PROTECTED DATA WITHIN LOCKING
MESSAGES USING A LOCK MANAGER
INDEPENDENT OF THE STORAGE
MECHANISM**

TECHNICAL FIELD

One embodiment of the invention relates to communica-
tions and computer systems employing locking mechanisms
to protect data; particularly, one embodiment relates to com-
municating protected data within locking messages; and
more particularly, one embodiment relates to bypassing
native storage by communicating protected data within lock-
ing messages with a lock manager independent of the storage
mechanism.

BACKGROUND

In many multiprocessor environments, inter-processor
communication is provided through shared global memory.
Access to this memory is generally protected through locks.
The latency for access to these resources is coupled through
the critical section of code and includes acquiring the lock,
reading the data, writing the data, and finally releasing the
lock. Note, nothing described or referenced in this document
is admitted as prior art to this application unless explicitly so
stated.

One such prior approach is illustrated in FIG. 1, which
shows how three requesters request and gain access to pro-
tected data. Note, as shown in FIG. 1, the lock manager is
independent of the storage mechanism as it does not access
the lock protected data from its native storage. In fact, the lock
manager of FIG. 1 never communicates nor otherwise pro-
cesses a value of the protected data.

As shown, each requester sends a request to the lock man-
ager, which provides independent access to the protected data
by sending a grant message to one of the requesters. The
granted requester then reads the protected data, performs its
processing, writes the protected data back to memory, and
then sends a release request to the lock manager to indicate
that it is done with the protected data. This process repeats and
a next requester is granted access to the data. Thus, there can
be significant amount of latency that is exposed in the critical
section of code, especially when multiple processors are
queued up behind a single locking queue. The significant
amount of latency is also true of processors that rely on
caching within the processor units to provide temporary stor-
age as well as provide direct inter-processor communication
to transfer data.

Also, one known system includes a lock manager in an I/O
subsystem which allows a message to include both locking
and data storage requests. This allows a requester proxy pro-
cess in the I/O subsystem to receive a message with both a
lock request and an I/O request. In response, this proxy pro-
cess makes a request to the lock manager for the lock, and in
response to a grant, it then makes the corresponding I/O
request corresponding to the I/O request and its native stor-
age. This approach may reduce some messages and latency
when the protected data is located in another subsystem, but
in response to each grant, the I/O native storage is still
accessed.

SUMMARY

Disclosed are, inter alia, methods, apparatus, data struc-
tures, computer-readable media, mechanisms, and means for

2

communicating protected data within locking messages. Data
is protected using locks, with the protected data sometimes
being included in the locking messages, which may reduce
overall processing latency, and/or reduce a bandwidth
requirement for and/or number of storage operations access-
ing the native storage of the protected data. For example, in
one embodiment, the lock manager receives lock requests
from each of the requesters, and selectively grants the lock
requests. The protected data is typically communicated in the
locking messages when the lock is highly contested, or at
least two request for access to the data are pending. The lock
manager initiates the sequence by indicating in a grant mes-
sage to a requester to include the protected data in its release
message. The lock manager then copies this data received in
the release message to its grant message to the next request-
er. Although this operation may require the lock manager to
temporarily store the received release message including the
protected data, it does not cache or otherwise store the pro-
tected data locally awaiting the receipt of a next request, for
example. If no other requesters are waiting, the grant message
includes an indication not to send the protected data, and thus
the requester typically stores this protected data to storage so
it can be accessed in the future.

One embodiment includes a lock manager configured to
control access via a lock to protected data maintained in
native storage independent of the lock manager. The lock
manager does not access the protected data from the native
storage, rather it copies the protected data received into grant
messages sent to a next requester. The lock manager is con-
figured to receive lock requests for the lock from multiple
requesters, and to selectively grant the lock requests which
includes communicating grants from the lock manager to the
plurality of requesters, wherein at least one of the communi-
cated grants includes the protected data.

In one embodiment, wherein at least one of the communi-
cated grants does not include the protected data. In one
embodiment, each of the communicated grants includes an
indication of whether or not the protected data is being com-
municated therewith. In one embodiment, each of the com-
municated grants includes an indication of whether or not the
protected data is requested to be sent to the lock manager with
a corresponding release of the lock. In one embodiment, each
of the lock requests includes an indication of whether or not
the corresponding one of the plurality of requesters will
accept the protected data from the lock manager.

One embodiment includes a lock manager that controls
access to protected data maintained in native storage inde-
pendent of the lock manager, wherein the lock manager does
not access the protected data from the native storage. The lock
manager receives a release of a lock for use in controlling
access to the protected data, with the received release includ-
ing the protected data. A next requester to be granted the lock
is identified in response to the receiving the release of the
lock. The protected data is copied from the release into a grant
message, and the grant message including the protected data
is sent to the next requester. In one embodiment, the grant
message includes an indication of that the protected data is
requested to be sent to the lock manager in a release message
corresponding to the grant message if another requester is
waiting for the lock, else an indication that the protected data
is not requested to be sent to the lock manager in the release
message.

One embodiment includes a lock manager that controls
access to protected data maintained in native storage inde-
pendent of the lock manager, wherein the lock manager does
not access the protected data from the native storage. The lock
manager receives locking requests for a lock controlling

③
Change to
"requester"

3

access to the protected data from a first requester and a second requester. The lock manager sends a first grant message not including the protected data to the first requester, and in response to identifying one or more requesters is waiting for the lock after the first requester, an indication to return the protected data is included in the grant message. A first release message including the protected data for the lock is subsequently received from the first requester.

In one embodiment, a second grant message is sent to the second requester, with the second grant message including the protected data and an indication of whether or not to send the protected data in a second release message. In one embodiment, the second grant message includes an indication to send the protected data in the second release message in response to identifying another requester is waiting for access to the lock. In one embodiment, the second grant message includes an indication not to send the protected data in the second release message in response to identifying another requester is not waiting for access to the lock. In one embodiment, the second grant message includes an indication not to send the protected data in the second release message, and in response to the indication not to send the protected data in the second release message, the second requester stores the protected data and does not include the protected data in the second release message.

BRIEF DESCRIPTION OF THE DRAWINGS

The appended claims set forth the features of the invention with particularity. The invention, together with its advantages, may be best understood from the following detailed description taken in conjunction with the accompanying drawings of which:

FIG. 1 illustrates a prior approach for using a lock to protect access to data;

FIG. 2 illustrates an approach used in one embodiment for protecting access to data with the protected data being communicated in conjunction with locking messages;

FIG. 3 illustrates locking messages used in one embodiment;

FIGS. 4A-B illustrate lock manager processes used in one embodiment;

FIG. 4C illustrates a requester process used in one embodiment; and

FIG. 5A illustrates a system including a lock manager and multiple requesters of one embodiment; and

FIG. 5B illustrate a system or component used in one embodiment for implementing a lock manager and/or one or more requesters.

DETAILED DESCRIPTION

Disclosed are, inter alia, methods, apparatus, data structures, computer-readable media, mechanisms, and means for communicating protected data within locking messages.

Embodiments described herein include various elements and limitations, with no one element or limitation contemplated as being a critical element or limitation. Each of the claims individually recites an aspect of the invention in its entirety. Moreover, some embodiments described may include, but are not limited to, inter alia, systems, networks, integrated circuit chips, embedded processors, ASICs, methods, and computer-readable media containing instructions. One or multiple systems, devices, components, etc. may comprise one or more embodiments, which may include some elements or limitations of a claim being performed by the same or different systems, devices, components, etc. The

4

embodiments described hereinafter embody various aspects and configurations within the scope and spirit of the invention, with the figures illustrating exemplary and non-limiting configurations.

As used herein, the term "packet" refers to packets of all types or any other units of information or data, including, but not limited to, fixed length cells and variable length packets, each of which may or may not be divisible into smaller packets or cells. The term "packet" as used herein also refers to both the packet itself or a packet indication, such as, but not limited to all or part of a packet or packet header, a data structure value, pointer or index, or any other part or direct or indirect identification of a packet or information associated therewith. For example, often times a router operates on one or more fields of a packet, especially the header, so the body of the packet is often stored in a separate memory while the packet header is manipulated, and based on the results of the processing of the packet (i.e., the packet header in this example), the entire packet is forwarded or dropped, etc. Additionally, these packets may contain one or more types of information, including, but not limited to, voice, data, video, and audio information. The term "item" is used generically herein to refer to a packet or any other unit or piece of information or data, a device, component, element, or any other entity. The phrases "processing a packet" and "packet processing" typically refer to performing some steps or actions based on the packet contents (e.g., packet header or other fields), and such steps or action may or may not include modifying, storing, dropping, and/or forwarding the packet and/or associated data.

The term "system" is used generically herein to describe any number of components, elements, sub-systems, devices, packet switch elements, packet switches, routers, networks, computer and/or communication devices or mechanisms, or combinations of components thereof. The term "computer" is used generically herein to describe any number of computers, including, but not limited to personal computers, embedded processing elements and systems, control logic, ASICs, chips, workstations, mainframes, etc. The term "processing element" is used generically herein to describe any type of processing mechanism or device, such as a processor, ASIC, field programmable gate array, computer, etc. The term "device" is used generically herein to describe any type of mechanism, including a computer or system or component thereof. The terms "task" and "process" are used generically herein to describe any type of running program, including, but not limited to a computer process, task, thread, executing application, operating system, user process, device driver, native code, machine or other language, etc., and can be interactive and/or non-interactive, executing locally and/or remotely, executing in foreground and/or background, executing in the user and/or operating system address spaces, a routine of a library and/or standalone application, and is not limited to any particular memory partitioning technique. The steps, connections, and processing of signals and information illustrated in the figures, including, but not limited to any block and flow diagrams and message sequence charts, may typically be performed in the same or in a different serial or parallel ordering and/or by different components and/or processes, threads, etc., and/or over different connections and be combined with other functions in other embodiments, unless this disables the embodiment or a sequence is explicitly or implicitly required (e.g., for a sequence of read the value, process the value—the value must be obtained prior to processing it, although some of the associated processing may be performed prior to, concurrently with, and/or after the read operation). Furthermore, the term "identify" is used generi-

(4) (5)
Change to
"requester"

cally to describe any manner or mechanism for directly or indirectly ascertaining something, which may include, but is not limited to receiving, retrieving from memory, determining, defining, calculating, generating, etc.

Moreover, the terms "network" and "communications mechanism" are used generically herein to describe one or more networks, communications media or communications systems, including, but not limited to the Internet, private or public telephone, cellular, wireless, satellite, cable, local area, metropolitan area and/or wide area networks, a cable, electrical connection, bus, etc., and internal communications mechanisms such as message passing, interprocess communications, shared memory, etc.

The term "message" is used generically herein to describe a piece of information which may or may not be, but is typically communicated via one or more communication mechanisms of any type.

The term "storage mechanism" includes any type of memory, storage device or other mechanism for maintaining instructions or data in any format. "Computer-readable medium" is an extensible term including any memory, storage device, storage mechanism, and other storage and signaling mechanisms including interfaces and devices such as network interface cards and buffers therein, as well as any communications devices and signals received and transmitted, and other current and evolving technologies that a computerized system can interpret, receive, and/or transmit. The term "memory" includes any random access memory (RAM), read only memory (ROM), flash memory, integrated circuits, and/or other memory components or elements. The term "storage device" includes any solid state storage media, disk drives, diskettes, networked services, tape drives, and other storage devices. Memories and storage devices may store computer-executable instructions to be executed by a processing element and/or control logic, and data which is manipulated by a processing element and/or control logic. The term "data structure" is an extensible term referring to any data element, variable, data structure, database, and/or one or more organizational schemes that can be applied to data to facilitate interpreting the data or performing operations on it, such as, but not limited to memory locations or devices, sets, queues, trees, heaps, lists, linked lists, arrays, tables, pointers, etc. A data structure is typically maintained in a storage mechanism. The terms "pointer" and "link" are used generically herein to identify some mechanism for referencing or identifying another element, component, or other entity, and these may include, but are not limited to a reference to a memory or other storage mechanism or location therein, an index in a data structure, a value, etc.

The term "one embodiment" is used herein to reference a particular embodiment, wherein each reference to "one embodiment" may refer to a different embodiment, and the use of the term repeatedly herein in describing associated features, elements and/or limitations does not establish a cumulative set of associated features, elements and/or limitations that each and every embodiment must include, although an embodiment typically may include all these features, elements and/or limitations. In addition, the phrase "means for xxx" typically includes computer-readable media containing computer-executable instructions for performing xxx.

In addition, the terms "first," "second," etc. are typically used herein to denote different units (e.g., a first element, a second element). The use of these terms herein does not necessarily connote an ordering such as one unit or event occurring or coming before another, but rather provides a mechanism to distinguish between particular units. Addition-

ally, the use of a singular tense of a noun is non-limiting, with its use typically including one or more of the particular thing rather than just one (e.g., the use of the word "memory" typically refers to one or more memories without having to specify "memory or memories," or "one or more memories" or "at least one memory", etc.). Moreover, the phrases "based on x" and "in response to x" are used to indicate a minimum set of items x from which something is derived or caused, wherein "x" is extensible and does not necessarily describe a complete list of items on which the operation is performed, etc. Additionally, the phrase "coupled to" is used to indicate some level of direct or indirect connection between two elements or devices, with the coupling device or devices modifying or not modifying the coupled signal or communicated information. The term "subset" is used to indicate a group of all or less than all of the elements of a set. The term "subtree" is used to indicate all or less than all of a tree. Moreover, the term "or" is used herein to identify a selection of one or more, including all, of the conjunctive items.

Locks can be used for many purposes. For example, one application of locks is described in Williams et al., "Using Ordered Locking Mechanisms to Maintain Sequences of Items Such as Packets," U.S. patent application Ser. No. 10/706,704, filed Nov. 12, 2003, which is hereby incorporated by reference.

Disclosed are, inter alia, methods, apparatus, data structures, computer-readable media, mechanisms, and means for communicating protected data within locking messages. Data is protected using locks, with the protected data sometimes being included in the locking messages, which may reduce overall processing latency, and/or reduce a bandwidth requirement for and/or number of storage operations accessing the native storage of the protected data. For example, in one embodiment, the lock manager receives lock requests from each of the requesters, and selectively grants the lock requests. The protected data is typically communicated in the locking messages when the lock is highly contested, or at least two request for access to the data are pending. The lock manager initiates the sequence by indicating in a grant message to a requester to include the protected data in its release message. The lock manager then copies this data received in the release message to its grant message to the next requester. Although this operation may require the lock manager to temporarily store the received release message including the protected data, it does not cache or otherwise store the protected data locally awaiting the receipt of a next request, for example. If no other requesters are waiting, the grant message includes an indication not to send the protected data, and thus the requester typically stores this protected data to storage so it can be accessed in the future.

One embodiment includes a lock manager configured to control access via a lock to protected data maintained in native storage independent of the lock manager. The lock manager does not access the protected data from the native storage, rather it copies the protected data received into grant messages sent to a next requester. The lock manager is configured to receive lock requests for the lock from multiple requesters, and to selectively grant the lock requests which includes communicating grants from the lock manager to the plurality of requesters, with at least one of the communicated grants includes the protected data.

In one embodiment, at least one of the communicated grants does not include the protected data. In one embodiment, each of the communicated grants includes an indication of whether or not the protected data is being communicated therewith. In one embodiment, each of the communicated grants includes an indication of whether or not the protected

⑥ ⑦
Change to
"requester"

7

data is requested to be sent to the lock manager with a corresponding release of the lock. In one embodiment, each of the lock requests includes an indication of whether or not the corresponding one of the plurality of requesters will accept the protected data from the lock manager.

One embodiment includes a lock manager that controls access to protected data maintained in native storage independent of the lock manager. The lock manager does not access the protected data from the native storage. The lock manager receives a release of a lock for use in controlling access to the protected data, with the received release including the protected data. A next requester to be granted the lock is identified in response to the receiving the release of the lock. The protected data is copied from the release into a grant message, and the grant message including the protected data is sent to the next requester. In one embodiment, the grant message includes an indication of that the protected data is requested to be sent to the lock manager in a release message corresponding to the grant message if another requester is waiting for the lock, else an indication that the protected data is not requested to be sent to the lock manager in the release message.

One embodiment includes a lock manager that controls access to protected data maintained in native storage independent of the lock manager. The lock manager does not access the protected data from the native storage. The lock manager receives locking requests for a lock controlling access to the protected data from a first requester and a second requester. The lock manager sends a first grant message not including the protected data to the first requester, and in response to identifying one or more requesters is waiting for the lock after the first requester, an indication to return the protected data is included in the grant message. A first release message including the protected data for the lock is subsequently received from the first requester.

In one embodiment, a second grant message is sent to the second requester, with the second grant message including the protected data and an indication of whether or not to send the protected data in a second release message. In one embodiment, the second grant message includes an indication to send the protected data in the second release message in response to identifying another requester is waiting for access to the lock. In one embodiment, the second grant message includes an indication not to send the protected data in the second release message in response to identifying another requester is not waiting for access to the lock. In one embodiment, the second grant message includes an indication not to send the protected data in the second release message, and in response to the indication not to send the protected data in the second release message, the second requester stores the protected data and does not include the protected data in the second release message.

One embodiment provides an indirect interprocess communication bypass channel through a lock mechanism that is connected to many processors. These processors normally communicate through shared global memory and use locks to enforce coherency. Under certain conditions, data can be transferred through lock messages instead of going through shared global memory. The data can be piggy-backed to the lock release message, go through the lock mechanism, and be piggy-backed to the lock grant message. The data is not stored for any significant amount of time in the lock mechanism. The lock messages typically include control signals to indicate when the conditions are right to use the bypass channel. This enforces the coherency of the shared memory location that might be bypassed.

8

When claiming to piggy-back the data to the lock message, the bypass channel could be either serial or parallel to the lock message channel, as long as there is a strong binding of lock messages to bypass data. In one embodiment, when requesting a lock, the request message includes an indication if it is willing to accept data through the bypass channel. When the lock is finally granted, the grant message indicates if it has data in the bypass channel, and if there is an entry following it in the locking queue that is willing to accept data through the bypass channel. If the grant indicates that data is present in the bypass channel, then the critical section can skip the read of the global shared memory location and use the data from the bypass channel instead.

If the grant indicates that the next entry in the locking queue is willing to accept data from the bypass channel, then the critical section of code can skip the write of the global shared memory location and send the data through the bypass channel instead. The critical section of code can always send the data through the bypass channel with the hope that a new arrival in the locking queue can use the data, but it must first commit the write to global shared memory if it is not certain. When the lock is released, an indication is made in the release message whether the bypass channel has data in it or not. The data in the bypass channel is typically not stored in memory in the lock mechanism; rather it is simply copied (possibly using a temporary storage location or register) from the release message and attached to the subsequent grant message.

Turning to the figures, FIG. 2 illustrates an approach used in one embodiment for protecting access to data with the protected data being communicated in conjunction with locking messages. Lock manager 200 receives locking requests 211-213 from requester-A 204, requester-B 206, and requester-C 208. Note, lock manager 200 is independent of the storage mechanism/protected data 202 as it does not access stored lock protected data 202 from its native storage as depicted in FIG. 2.

An example of such a locking request is locking request message 300 illustrated in FIG. 3. As shown, request message 300 includes an indication 301 of which lock is being requested, an identification 302 of the requester, and an indication 303 of whether the requester supports protected data in locking messages. Of course, one embodiment uses another communication mechanism and/or some, all, or none of values 301-303. Also, the number of bits illustrated for certain fields are only exemplary of one embodiment.

For purposes of illustration, the description of FIG. 2 will assume that all requesters will always support protected data in locking messages (and thus indication 303 is not required). In one embodiment when all requesters do not always support protected data being communicated in locking messages, locking manager 200 will only request protected data when the next requester supports such, and provide protected data in a locking message to a supporting requester.

Also shown in FIG. 3, is a grant message 310 used in one embodiment, and release message 320 used in one embodiment. As shown, grant message 310 includes a lock indication 311 identifying the lock, a field 312 for including protected data, an indication 313 of whether or not field 312 is populated with the protected data, and an indication 314 of whether or not an explicit request to bypass native storage for the protected data is being made.

In one embodiment, if the native storage bypass request indication is set, the requester must return the latest value of the protected data in a corresponding release message, and the requester may or may not store the protected data in its native storage prior to sending the release (i.e., the value of the protected data in the native storage may or may not be the

(8)
change to
"requester"

11

block 458, the protected data is retrieved from storage, and processing is performed based on the retrieved protected data.

After processing of the protected data is complete, then, as determined in process block 460, if the received grant indicated that native storage is to be bypassed, then processing proceeds to process block 468. Otherwise, in process block 462, the latest value of the protected data is stored in its native storage location. As determined in process block 464, if the protected data is to be included in the release (e.g., it is always included, or optionally included based on some determination), then processing proceeds to process block 468. Otherwise, in process block 466, the release message is sent to the lock manager without the protected data. In process block 468, a release message including the latest value of the protected data is sent to the lock manager. Processing is complete as indicated by process block 469.

FIG. 5A illustrates a system including a lock manager 501 and multiple requesters 511-519 of one embodiment. FIG. 5A shows the extensible nature of one embodiment of the invention which can be applied to an application. Lock manager 501 and multiple requesters 511-519 can be processes, separate processing elements, or any other processing mechanism or entity in one or more systems, elements, or components. As shown, lock manager 501 is communicatively coupled via communications mechanism 509 with multiple requesters 511-519, which are also communicatively coupled to the mechanism 502 for storing the protected data and/or other resources.

FIG. 5B illustrate a system or component used in one embodiment for implementing a lock manager and/or one or more requesters. In one embodiment, system or component 540 performs one or more processes corresponding to one of the flow diagrams illustrated or otherwise described herein. For example, in one embodiment, the lock manager and requesters are processes running on processing element 541, and memory 542 is used for storing the protected data when it is not communicated via locking messages.

In one embodiment, system or component 540 includes a processing element 541, memory 542, storage devices 543, and an interface 544 for sending and receiving packets, items, and/or other information, which are typically coupled via one or more communications mechanisms 549 (shown as a bus for illustrative purposes.) Various embodiments of component 540 may include more or less elements. The operation of component 540 is typically controlled by processing element 541 using memory 542 and storage devices 543 to perform one or more tasks or processes. Memory 542 is one type of computer-readable media, and typically comprises random access memory (RAM), read only memory (ROM), flash memory, integrated circuits, and/or other memory components. Memory 542 typically stores computer-executable instructions to be executed by processing element 541 and/or data which is manipulated by processing element 541 for implementing functionality in accordance with an embodiment. Storage devices 543 are another type of computer-readable media, and typically comprise solid state storage media, disk drives, diskettes, networked services, tape drives, and other storage devices. Storage devices 543 typically store computer-executable instructions to be executed by processing element 541 and/or data which is manipulated by processing element 541 for implementing functionality in accordance with an embodiment.

In view of the many possible embodiments to which the principles of our invention may be applied, it will be appreciated that the embodiments and aspects thereof described herein with respect to the drawings/figures are only illustrative and should not be taken as limiting the scope of the

12

invention. For example and as would be apparent to one skilled in the art, many of the process block operations can be re-ordered to be performed before, after, or substantially concurrent with other operations. Also, many different forms of data structures could be used in various embodiments. The invention as described herein contemplates all such embodiments as may come within the scope of the following claims and equivalents thereof.

What is claimed is:

1. An apparatus for protecting data using locks, the apparatus comprising: one or more processors and memory, configured to include:

a lock manager configured to control access via a lock to protected data maintained in native storage independent of the lock manager, wherein the lock manager does not access said protected data from said native storage; and a plurality of requesters;

wherein the lock manager is configured to receive lock requests for the lock from each of the plurality of requesters, and to selectively grant said lock requests which includes communicating grants from the lock manager to the plurality of requesters, wherein at least one of said communicated grants includes said protected data received in a corresponding release of the lock message from a previous holder of the lock of the plurality of requesters.

2. The apparatus of claim 1, wherein at least one of said communicated grants does not include said protected data.

3. The apparatus of claim 1, wherein each of said communicated grants includes an indication of whether or not said protected data is being communicated therewith.

4. The apparatus of claim 1, wherein each of said communicated grants includes an indication of whether or not said protected data is requested to be sent to the lock manager with a corresponding release of the lock.

5. The apparatus of claim 1, wherein each of said lock requests includes an indication of whether or not the corresponding one of the plurality of requesters will accept said protected data from the lock manager.

6. A method performed by a lock manager controlling access to protected data maintained in native storage independent of the lock manager, wherein the lock manager does not access said protected data from said native storage, the method comprising:

receiving a release of a lock for use in controlling access to said protected data, the release including said protected data;

identifying a next requester to be granted the lock in response to said receiving the release of the lock;

copying said protected data from the release into a grant message; and

sending the grant message to the next requester, the grant message including said protected data.

7. The method of claim 6, wherein the grant message includes an indication of that said protected data is requested to be sent to the lock manager in a release message corresponding to the grant message if another requester is waiting for the lock, else an indication that said protected data is not requested to be sent to the lock manager in the release message.

8. A tangible computer-readable medium storing thereon computer-executable instructions for performing steps by a lock manager for controlling access to protected data maintained in native storage independent of the lock manager, wherein the lock manager does not access said protected data from said native storage, said steps comprising:

⑨ change to "requesters"

13

receiving a release of a lock for use in controlling access to said protected data, the release including said protected data;

identifying a next requester to be granted the lock in response to said receiving the release of the lock;

copying said protected data from the release into a grant message; and

sending the grant message to the next requester, the grant message including said protected data.

9. The computer-readable medium of claim 8, wherein the grant message includes an indication of that said protected data is requested to be sent to the lock manager in a release message corresponding to the grant message if another requester is waiting for the lock, else an indication that said protected data is not requested to be sent to the lock manager in the release message.

10. A lock manager controlling access to protected data maintained in native storage independent of the lock manager, wherein the lock manager does not access said protected data from said native storage, the lock manager comprising:

means for receiving a release of a lock for use in controlling access to said protected data, the release including said protected data;

means for identifying a next requester to be granted the lock in response to said receiving the release of the lock;

means for copying said protected data from the release into a grant message and for sending the grant message to the next requester.

11. The lock manager of claim 10, means for including in the grant message an indication of that said protected data is requested to be sent to the lock manager in a release message corresponding to the grant message if another requester is waiting for the lock, else an indication that said protected data is not requested to be sent to the lock manager in the release message.

12. A method performed by a lock manager controlling access to protected data maintained in native storage independent of the lock manager, wherein the lock manager does not access said protected data from said native storage, the method comprising:

receiving locking requests for a lock controlling access to said protected data from a first requester and a second requester;

sending a first grant message to the first requester, the first grant message not including said protected data, and in response to identifying one or more requesters is waiting for the lock after the first requester, including an indication to return said protected data in the first grant message;

receiving a first release message corresponding to the first grant message for the lock from the first requester, the first release message including said protected data and sending a second grant message to the second requester, the second grant message including said protected data received in the first release message.

13. The method of claim 12, wherein the second grant message includes an indication of whether or not to send said protected data in a second release message.

14. The method of claim 13, wherein the second grant message includes an indication to send said protected data in the second release message in response to identifying another requester's waiting for access to the lock.

15. The method of claim 13, wherein the second grant message includes an indication not to send said protected data in the second release message in response to identifying another requester's not waiting for access to the lock.

14

16. The method of claim 13, wherein the second grant message includes an indication not to send said protected data in the second release message; and the method comprises in response to said indication not to send said protected data in the second release message, the second requester storing said protected data and not including said protected data in the second release message.

17. A tangible computer-readable medium storing thereon computer-executable instructions for performing steps by a lock manager for controlling access to protected data maintained in native storage independent of the lock manager, wherein the lock manager does not access said protected data from said native storage, said steps comprising:

receiving locking requests for a lock controlling access to said protected data from a first requester and a second requester;

sending a first grant message to the first requester, the first grant message not including said protected data, and in response to identifying one or more requesters is waiting for the lock after the first requester, including an indication to return said protected data in the first grant message;

receiving a first release message corresponding to the first grant message for the lock from the first requester, the first release message including said protected data; and sending a second grant message to the second requester, the second grant message including said protected data received in the first release message.

18. The computer-readable medium of claim 17, wherein the second grant message includes an indication of whether or not to send said protected data in a second release message.

19. The computer-readable medium of claim 18, wherein the second grant message includes an indication to send said protected data in the second release message in response to identifying another requester's waiting for access to the lock.

20. The computer-readable medium of claim 18, wherein the second grant message includes an indication not to send said protected data in the second release message in response to identifying another requester's not waiting for access to the lock.

21. The computer-readable medium of claim 18, wherein the second grant message includes an indication not to send said protected data in the second release message; and said steps comprise in response to said indication not to send said protected data in the second release message, the second requester storing said protected data and not including said protected data in the second release message.

22. A lock manager controlling access to protected data maintained in native storage independent of the lock manager, wherein the lock manager does not access said protected data from said native storage, the lock manager comprising:

means for receiving locking requests for a lock controlling access to said protected data from a first requester and a second requester;

means for sending a first grant message to the first requester, the first grant message not including said protected data, and in response to identifying one or more requesters is waiting for the lock after the first requester, including an indication to return said protected data in the first grant message;

(10) (11) (12)
change to
"requester"

(13) (14)
change to
"requester"

15

means for receiving a first release message for the lock from the first requester, the first release message including said protected data; and means for sending a second grant message to the second requester, the second grant message including said protected data received in the first release message.

23. The lock manager of claim 22, wherein the second grant message includes an indication of whether or not to send said protected data in a second release message.

24. The lock manager of claim 23, comprising means for including in the second grant message an indication to send said protected data in the second release message in response to identifying another requester is waiting for access to the lock.

16

25. The lock manager of claim 23, comprising means for including in the second grant message an indication not to send said protected data in the second release message in response to identifying another requester is not waiting for access to the lock.

26. The lock manager of claim 23, comprising: means for including in the second grant message an indication not to send said protected data in the second release message; and means for the second requester to store said protected data and not to include said protected data in the second release message in response to said indication not to send said protected data in the second release message.

* * * * *

16

change to
"requester"

15

change
"requester"